



(19) 대한민국특허청(KR)  
(12) 공개특허공보(A)

(11) 공개번호 10-2015-0008342

(43) 공개일자 2015년01월22일

(51) 국제특허분류(Int. Cl.)

G06Q 50/10 (2012.01)

(21) 출원번호 10-2014-0087544

(22) 출원일자 2014년07월11일

심사청구일자 없음

(30) 우선권주장

13306002.0 2013년07월12일

유럽특허청(EPO)(EP)

(71) 출원인

툼슨 라이센싱

프랑스 92130 이씨레물리노 잔 다르크 뒤편 1-5

(72) 발명자

라나경, 제임스

프랑스 35576 쉐송 쉐비네 쉐에스 176 16 자크 데  
상 블랑 아브뉴 데 상 블랑 975 페르니폴로르 에  
르 에 데 프랑스

계경, 마리

프랑스 35576 쉐송 쉐비네 쉐에스 176 16 자크 데  
상 블랑 아브뉴 데 상 블랑 975 페르니폴로르 에  
르 에 데 프랑스

슈무게르, 필립

프랑스 35576 쉐송 쉐비네 쉐에스 176 16 자크 데  
상 블랑 아브뉴 데 상 블랑 975 페르니폴로르 에  
르 에 데 프랑스

(74) 대리인

양영준, 전경석, 백만기

전체 청구항 수 : 총 12 항

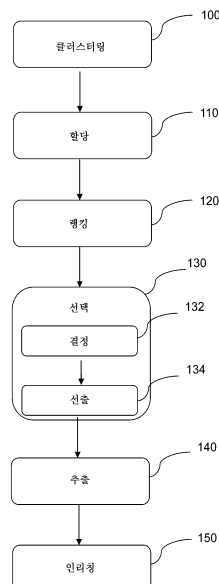
(54) 발명의 명칭 멀티미디어 콘텐츠를 인리칭하기 위한 방법 및 대응하는 디바이스

### (57) 요약

멀티미디어 콘텐츠를 인리칭하기 위한 방법, 및 대응하는 디바이스이다. 본 발명에 따르면, 방법은: 관련 크리에이터들이라고 명명되는 적어도 2개의 크리에이터들을 포함하는 클러스터들을 전달하기 위해 상기 복수의 크리에이터들을 클러스터링하는 단계(100); 제1 클러스터라고 명명되는 제1 크리에이터를 포함하는 클러스터의 적어

(뒷면에 계속)

대표도 - 도1



도 일부 관련 크리에이터들에, 상기 관련 크리에이터들과 상기 제1 크리에이터 사이의 유사성의 적어도 하나의 기준으로부터 결정된 관련성 스코어를 할당하는 단계(110); 그들의 관련성 스코어들에 따라 상기 제1 클러스터의 관련 크리에이터들을 랭킹시키는 단계(120); 복수의 멀티미디어 콘텐츠들 중에서, 상기 제1 멀티미디어 콘텐츠에 관련되며 상기 제1 클러스터의 랭킹된 관련 크리에이터에 의해 전달되는 적어도 하나의 제2 멀티미디어 콘텐츠를 선택하는 단계(130); 상기 제2 멀티미디어 콘텐츠에 기초하여 상기 제1 멀티미디어 콘텐츠를 인리칭하는 단계(150)를 포함한다.

---

## 특허청구의 범위

### 청구항 1

복수의 멀티미디어 콘텐츠 크리에이터(creator)(202, 204, 206, 208, 212, 214, 222, 224, 226) 중에서, 제1 크리에이터(204)에 의해 생성된 제1 멀티미디어 콘텐츠를 인리칭(enriching)하기 위한 방법으로서,

관련 크리에이터들이라고 명명되는 적어도 2개의 크리에이터들(202, 204, 206, 208)을 포함하는 클러스터들(200, 210, 220)을 전달하기 위해 상기 복수의 크리에이터들을 클러스터링하는 단계(100);

제1 클러스터(200)라고 명명되는, 상기 제1 크리에이터(204)를 포함하는 상기 클러스터의 적어도 일부 관련 크리에이터들에, 상기 관련 크리에이터들(202, 206, 208)과 상기 제1 크리에이터(204) 사이의 유사성의 적어도 하나의 기준으로부터 결정된 관련성 스코어(232, 234, 236)를 할당하는 단계(110);

그들의 관련성 스코어들에 따라 상기 제1 클러스터(200)의 관련 크리에이터들(202, 206, 208)을 랭킹(ranking)시키는 단계(120);

복수의 멀티미디어 콘텐츠들 중에서, 상기 제1 멀티미디어 콘텐츠에 관련되며, 상기 제1 클러스터(200)의 랭킹된 관련 크리에이터(202, 206, 208)에 의해 전달되는 적어도 하나의 제2 멀티미디어 콘텐츠를 선택하는 단계(130); 및

상기 제2 멀티미디어 콘텐츠에 기초하여 상기 제1 멀티미디어 콘텐츠를 인리칭하는 단계(150)

를 포함하는 것을 특징으로 하는 방법.

### 청구항 2

제1항에 있어서,

상기 제2 멀티미디어 콘텐츠로부터, 상기 제1 멀티미디어 콘텐츠에 관련된 적어도 하나의 엘리먼트를 추출하는 추가의 단계(140)를 포함하는 방법.

### 청구항 3

제2항에 있어서,

상기 추출하는 단계는 자연어(natural language) 프로세싱에 의존하는 방법.

### 청구항 4

제1항 내지 제3항 중 어느 한 항에 있어서,

상기 클러스터링하는 단계는 식별된 특징들의 세트에 기초하여 K-평균 알고리즘(K-means algorithm)을 구현하는 방법.

### 청구항 5

제1항 내지 제4항 중 어느 한 항에 있어서,

상기 선택된 제2 멀티미디어 콘텐츠는 할당된 관련성 스코어들 중에서 가장 높은 관련성 스코어를 가지는 상기 제1 클러스터(200)의 랭킹된 관련 크리에이터(202, 206, 208)에 의해 생성되는 방법.

### 청구항 6

제1항 내지 제5항 중 어느 한 항에 있어서,

상기 제1 멀티미디어 콘텐츠에 관련되며, 관련성 임계치와 적어도 동일한 관련성 스코어를 가지는 상기 제1 클러스터(200)의 랭킹된 관련 크리에이터들(202, 206, 208)에 의해 전달되는 몇몇 제2 멀티미디어 콘텐츠들의 선택을 포함하는 방법.

#### 청구항 7

제1항 내지 제6항 중 어느 한 항에 있어서,  
상기 유사성의 기준은 거리 기준인 방법.

#### 청구항 8

제1항 내지 제7항 중 어느 한 항에 있어서,  
상기 제1 멀티미디어 콘텐츠는 텍스트 콘텐츠인 방법.

#### 청구항 9

복수의 멀티미디어 콘텐츠 크리에이터(202, 204, 206, 208, 212, 214, 222, 224, 226) 중에서, 제1 크리에이터에 의해 생성된 제1 멀티미디어 콘텐츠를 인리칭하기 위한 디바이스로서,  
관련 크리에이터들이라고 명명되는 적어도 2개의 크리에이터들(202, 204, 206, 208)을 포함하는 클러스터들(200, 210, 220)을 전달하기 위해 상기 복수의 크리에이터들을 클러스터링하도록 구성된 클러스터링 모듈(112);  
제1 클러스터라고 명명되는, 상기 제1 크리에이터(204)를 포함하는 상기 클러스터(200)의 각각의 관련 크리에이터들에, 상기 관련 크리에이터와 상기 제1 크리에이터 사이의 유사성의 적어도 하나의 기준으로부터 결정된 관련성 스코어를 할당하도록 구성된 할당 모듈(106);  
그들의 관련성 스코어들에 따라 상기 제1 클러스터의 관련 크리에이터들을 랭킹시키도록 구성된 랭킹 모듈(104);  
복수의 멀티미디어 콘텐츠들 중에서, 상기 제1 멀티미디어 콘텐츠에 관련되며, 상기 제1 클러스터의 랭킹된 관련 크리에이터에 의해 전달되는 적어도 하나의 제2 멀티미디어 콘텐츠를 선택하도록 구성된 선택 모듈(131); 및  
상기 제2 멀티미디어 콘텐츠에 기초하여 상기 제1 멀티미디어 콘텐츠를 인리칭하도록 구성된 인리칭 모듈(108)을 포함하는 것을 특징으로 하는 디바이스.

#### 청구항 10

제9항에 있어서,  
상기 제2 멀티미디어 콘텐츠로부터, 상기 제1 멀티미디어 콘텐츠에 관련된 적어도 하나의 엘리먼트를 추출하도록 구성된 추출 모듈(133)을 더 포함하는 디바이스.

#### 청구항 11

통신 네트워크로부터 다운로드 가능한 것, 그리고 컴퓨터에 의해 판독 가능하고 그리고 프로세서에 의해 실행 가능한 것 중 적어도 하나인 매체 상에 레코딩되는 것 중 적어도 하나인 컴퓨터 프로그램 제품으로서, 제1항 내지 제8항 중 적어도 한 항에 따른 방법의 단계들을 구현하기 위한 프로그램 코드 명령들을 포함하는 컴퓨터 프로그램 제품.

#### 청구항 12

레코딩된 컴퓨터 프로그램 제품을 포함하며 프로세서에 의해 실행될 수 있는 비-일시적 컴퓨터-판독가능한 매체로서, 제1항 내지 제8항 중 적어도 한 항에 따른 방법의 단계들을 구현하기 위한 프로그램 코드 명령들을 포함하는 비-일시적 컴퓨터-판독가능한 매체.

### 명세서

#### 기술분야

발명의 분야는 일반적으로, 예를 들어, 네트워크(예를 들어, 인터넷)를 통해 발행되는 멀티미디어 콘텐츠에 관한 것이고, 더 구체적으로는 메타데이터(예를 들어, 기록된 코멘터리, 오디오 콘텐츠, 비디오 클립 등)를 이용하여 이러한 멀티미디어 콘텐츠를 인리칭하기 위한 기법에 관한 것이다.

[0001]

## 배경 기술

- [0002] 추가 정보(소위 메타데이터)를 가지는 멀티미디어 콘텐츠(예컨대, 오디오/비디오 콘텐츠 또는 기록된 기사)의 주석은 엄청난 양의 이용가능한 데이터로 인해, 더욱더 대처해야 할 주요 이슈가 되고 있다. 이러한 메타데이터는 전문 소스로부터 올 수 있지만, 또한 소셜 네트워크 및 웹 포럼과 같은 유저 소스로부터 올 수도 있다.
- [0003] 이러한 엄청난 양의 데이터를 대하기 위해, 주어진 멀티미디어 콘텐츠에 대한 가장 많이 관련된 추가 정보를 식별하고 필터링하기 위한 필요성이 존재하며, 이러한 정보 피스들은 예를 들어, 인터넷, 웹 포럼, 전용 웹사이트 등을 통해 소셜 네트워크로부터 검색된다.
- [0004] 종래 기술은 인터넷을 통해 주어진 멀티미디어 콘텐츠에 연관된 메타데이터를 탐색하고 액세스하기 위해, 종종 키워드 또는 고정된 분류를 사용하여 검색 엔진을 구현한다. 그럼에도 불구하고, 현재 검색 엔진은 요청 유저들에게 전달하기 이전에 메타데이터의 관련성을 고려하지 않으며, 다만 주어진 멀티미디어 콘텐츠 내에 있거나 주어진 멀티미디어 콘텐츠와 연관된 요청된 키워드를 단지 사용한다.
- [0005] 추가로, 온라인으로 발행된 주어진 멀티미디어 콘텐츠에 대해, 예를 들어, "더하기" 또는 "유사한" 동작으로 인해 대다수의 웹 유저에 의해 인증된 관련 메타데이터를 선택하는 것이 알려져 있다. 그러나, 이러한 관련 메타데이터의 식별은 다른 웹 유저의 상호작용을 요구한다.
- [0006] 따라서, 주어진 멀티미디어 콘텐츠를 인리칭하기 위해 가장 관심있는 메타데이터를 식별하는 것에 대한 상당한 요구가 존재한다.

## 발명의 내용

- [0007] 따라서, 발명의 목적은 이전에 언급된 결점들의 적어도 일부를 처리하는 것이며, 특히, 주어진 멀티미디어 콘텐츠에 가깝게 링크된 관련 추가 정보의 검색을 개선하는 것이다.
- [0008] 따라서, 적어도 하나의 실시예에 따르면, 본 발명의 해법은 멀티미디어 콘텐츠를 인리칭하는 새로운 방식을 제공한다.
- [0009] 전술된 문제점은, 게시내용의 일 양상에 따라, 복수의 멀티미디어 콘텐츠 크리에이터 중, 제1 크리에이터에 의해 생성된 제1 멀티미디어 콘텐츠를 인리칭하기 위한 방법에 의해 해결될 수 있다.
- [0010] 발명에 따르면, 이러한 방법은 후속하는 단계:
- [0011] - 관련 크리에이터라고 명명된 적어도 2개의 크리에이터를 포함하는 클러스터를 전달하기 위해 상기 복수의 크리에이터를 클러스터링하는 단계;
- [0012] - 제1 클러스터라고 명명된, 제1 크리에이터를 포함하는 클러스터의 적어도 일부 관련 크리에이터에, 상기 관련 크리에이터와 상기 제1 크리에이터 사이의 유사성의 적어도 하나의 기준으로부터 결정된 관련성 스코어를 할당하는 단계;
- [0013] - 그들의 관련성 스코어에 따라 상기 제1 클러스터의 관련 크리에이터를 랭킹시키는 단계;
- [0014] - 복수의 멀티미디어 콘텐츠 중에서, 상기 제1 멀티미디어 콘텐츠와 관련되며 상기 제1 클러스터의 랭킹된 관련 크리에이터에 의해 전달된 적어도 하나의 제2 멀티미디어 콘텐츠를 선택하는 단계;
- [0015] - 상기 제2 멀티미디어 콘텐츠에 기초하여 상기 제1 멀티미디어 콘텐츠를 인리칭하는 단계
- [0016] 를 포함한다.
- [0017] 따라서, 본 발명은 가장 큰 잠재 값을 가지는 추가 데이터를 자동으로 전달하기 위해 주어진 멀티미디어 콘텐츠에 관련된 이용가능한 멀티미디어 데이터에 적용되는 필터로서 작용할 수 있다. 이후, 주어진 콘텐츠 및 추가 콘텐츠 모두의 크리에이터들 사이의 유사성(다시 말해, 이들이 서로 얼마나 가까이 관련되는지)을 고려함으로써, (예를 들어, 인터넷으로부터) 자동으로 검색되고 관련된 것으로서 고려된 하나 이상의 멀티미디어 콘텐츠로 인해 상기 주어진 멀티미디어 콘텐츠를 인리칭하는 것이 허용될 수 있다. 이는 주어진 멀티미디어 콘텐츠의 더 풍부한 디스크립션(description)을 허용한다.
- [0018] 바람직한 실시예에 따르면, 상기 방법은, 제2 멀티미디어 콘텐츠로부터, 상기 제1 멀티미디어 콘텐츠에 관련된 적어도 하나의 엘리먼트를 추출하는 추가의 단계를 포함한다.

- [0019] 추가로, 상기 추출하는 단계는 자연어(natural language) 프로세싱에 의존할 수 있다.
- [0020] 또한, 상기 클러스터링 단계는 식별된 특징들에 세트에 기초하여 K-평균 알고리즘(K-means algorithm)을 실행할 수 있다. 명백하게, 다른 종류의 클러스터링 알고리즘이 사용될 수 있다.
- [0021] 크리에이터와 연관된 특징은 특히, 관심 있는 도메인, 언어 스타일, 크리에이터의 활동의 일부 소셜 표시자 등을 포함할 수 있다. 일부 특징은, 예를 들어, 엔티티 인식(예를 들어, 특정 위치의, 문자적 명칭의 인식 등), 또는 특징-기반 정서 분석(또한 "오피니언 마이닝(opinion mining)"이라고 명명됨)이라고 명명된, 정보 추출과 같은 자연어 프로세싱 기법을 통해, 온라인 커뮤니티 내의 크리에이터의 프로파일로부터 자동으로 추출될 수 있다. 이러한 특징은 예를 들어, "전문가", "비전문가", "로봇" 등과 같은, 크리에이터의 자격을 포함할 수 있다. 일부 다른 특징은, 그래프-기반 또는 수치적 측정 기법을 이용하여, 온라인 커뮤니티(예를 들어, 영화광의 포럼) 내의 크리에이터 소셜 중심성 및 생산성을 분석함으로써 수집될 수 있다. 이러한 특징은, 발행된 다수의 멀티미디어 콘텐츠, 또는 스타일리스트 언어 디스크립터, 또는 생성된 콘텐츠에 대해 추출된 엘리먼트(예를 들어, 크리에이터 콘텐츠 내에 빈번하게 존재하는 엘리먼트)와 같은, 크리에이터의 기여도와 관련될 수 있다.
- [0022] 각각의 클러스터는 가까운 특징 세트를 가지는 크리에이터(특징들의 세트가 유사성의 기준에 따라 이들에 대한 가까운 값을 획득하게 되는 자들이라고도 함)를 포함할 수 있다.
- [0023] 추가로, 상기 선택된 제2 멀티미디어 콘텐츠는 할당된 관련성 스코어 중에서 가장 높은 관련성 스코어를 가지는 상기 제1 클러스터의 랭킹된 관련 크리에이터에 의해 생성될 수 있다.
- [0024] 또한, 상기 방법은 상기 제1 멀티미디어 콘텐츠와 관련되며, 관련성 임계와 적어도 동일한 관련성 스코어를 가지는 상기 제1 클러스터의 랭킹된 관련 크리에이터에 의해 전달되는 몇몇 제2 멀티미디어 콘텐츠의 선택을 포함할 수 있다.
- [0025] 또한, 상기 유사성의 기준은 유리하게는 거리 기준(예컨대, 유클리드 거리, 맨하탄 거리 등)일 수 있다.
- [0026] 예로서, 상기 제1 멀티미디어 콘텐츠는 텍스트 콘텐츠일 수 있다. 이러한 텍스트 콘텐츠는 예를 들어, 제품, 서비스 또는 이벤트 코멘트, 이야기, 블로그 또는 포럼 포스트와 같은 웹 유저에 의해 기록된 임의의 종류의 텍스트를 나타낼 수 있다.
- [0027] 게시내용은 추가로, 복수의 멀티미디어 콘텐츠 크리에이터 중, 제1 크리에이터에 의해 생성된 제1 멀티미디어 콘텐츠를 인리칭하기 위한 디바이스에 관련된다.
- [0028] 발명에 따르면, 상기 디바이스는 후속하는 모듈:
- [0029] - 관련 크리에이터라고 명명된, 적어도 2개의 크리에이터를 포함하는 클러스터를 전달하기 위해 상기 복수의 크리에이터를 클러스터링하도록 구성되는 클러스터링 모듈;
- [0030] - 제1 클러스터라고 명명된 제1 크리에이터를 포함하는 클러스터의 각각의 관련된 크리에이터에, 상기 관련 크리에이터와 상기 제1 크리에이터 사이의 유사성의 적어도 하나의 기준으로부터 결정된 관련성 스코어를 할당하도록 구성되는 할당 모듈;
- [0031] - 그들의 관련성 스코어에 따라 제1 클러스터의 관련된 크리에이터를 랭킹시키도록 구성되는 랭킹 모듈;
- [0032] - 복수의 멀티미디어 콘텐츠 중에서, 제1 멀티미디어 콘텐츠와 관련되며 제1 클러스터의 랭킹된 관련 크리에이터에 의해 전달되는 적어도 하나의 제2 멀티미디어 콘텐츠를 선택하도록 구성되는 선택 모듈;
- [0033] - 상기 제2 멀티미디어 콘텐츠에 기초하여 상기 제1 멀티미디어 콘텐츠를 인리칭하도록 구성되는 인리칭 모듈
- [0034] 을 포함한다.
- [0035] 특히, 상기 디바이스는, 상기 제2 멀티미디어 콘텐츠로부터, 상기 제1 멀티미디어 콘텐츠에 관련된 적어도 하나의 엘리먼트를 추출하도록 적응되는 추출 모듈을 더 포함한다.
- [0036] 게시내용은 추가로, 멀티미디어 콘텐츠를 인리칭하기 위한 상기 방법을 구현하기 위한 프로그램 코드 명령을 포함하는, 통신 네트워크로부터 다운로드가능한 것, 그리고 컴퓨터에 의해 판독가능하고 그리고 프로세서에 의해 실행가능한 것 중 적어도 하나인 매체 상에 레코딩된 것 중 적어도 하나인 컴퓨터 프로그램 제품에 관련된다.
- [0037] 게시 내용은 추가로, 멀티미디어 콘텐츠를 인리칭하기 위한 상기 방법을 구현하기 위한 프로그램 코드 명령을

포함하는, 레코딩된 컴퓨터 프로그램 제품을 포함하고 프로세서에 의해 실행될 수 있는 비-일시적 컴퓨터-판독 가능한 매체에 관련된다.

[0038] 발명의 예들이 위에서 제한된 개수의 실시예들에 관련하여 기술되었지만, 당업자는, 이 기재를 읽을 때, 이 발명의 범위를 이탈하지 않고 다른 실시예들이 연상될 수 있음을 이해할 것이다.

[0039] 특히, 명시적으로 기술되어 있지는 않지만, 본 실시예들은 임의의 조합 또는 세부조합으로 사용될 수 있다.

### 도면의 간단한 설명

[0040] 발명의 실시예의 다른 특성 및 장점은, 비-제한적이며 지시적인 예(발명의 모든 실시예는 하기에 기술되는 실시예들의 특성들 및 장점들에 제한되지 않음)에 의해 제공되는, 후속하는 기재 및 첨부 도면을 읽을 때 명백해질 것이다.

도 1은 본 발명의 바람직한 실시예에 따른 방법의 주요 단계들을 도시하는 도면이다.

도 2는 멀티미디어 콘텐츠의 복수의 크리에이터를 나타내는 도면이다.

도 3은 도 1에 기술되어 있는 방법을 구현하도록 적응되는 디바이스의 예의 블록도이다.

도 3은 물리적으로 분리된 엔티티들에 반드시 대응하지는 않는 순수한 기능 엔티티들이다. 소위, 이들은 소프트웨어, 하드웨어의 형태로 개발되거나, 또는 하나 이상의 프로세서를 포함하는 하나 또는 몇몇 집적 회로에서 구현될 수 있다.

가능할 때마다, 동일한 참조 번호는 도면 전반에 걸쳐 동일하거나 유사한 부분들을 지칭하도록 사용될 것이다.

### 발명을 실시하기 위한 구체적인 내용

[0041] 발명의 바람직한 실시예에 따르면, 도 1에 도시되어 있는 바와 같이, 고려된 콘텐츠에 관련되며 다른 크리에이터에 의해 네트워크(예컨대, 인터넷) 상에서 발행된 추가적인 멀티미디어 콘텐츠의 검색 및 분석으로 인해 - 주어진 크리에이터에 의해 구축된 - 고려된 멀티미디어 콘텐츠를 인리칭하기 위한 새로운 창의적인 방법 M이 제안된다. 추가적인 멀티미디어 콘텐츠는 다른 크리에이터에 의해 생성되었을 수 있다.

[0042] 멀티미디어 콘텐츠가 영화 리뷰, A/V 콘텐츠(비디오 클립으로서), 이야기, 리포트(예를 들어, 스포츠 경기 리포트)등과 같이 다양할 수 있다는 점이 이해되어야 한다.

[0043] 또한, 도 2에 도시되어 있는 바와 같이, 멀티미디어 콘텐츠 크리에이터(202, 204, 206, 208, 212, 214, 222, 224, 226)는, 예를 들어, 영화에 전용되는 웹 포럼, 소셜 네트워크 등과 같은 온라인 커뮤니티의 일부 또는 모든 멤버들에 대응할 수 있다. 상기 복수의 크리에이터는 멀티미디어 기사(예컨대, 웹 신문에 대한 기록된 기사, 사진 또는 비디오)를 생성하거나 전달하는 전문가, 및 전문가에 의해 생성된 이벤트, 제품, 서비스, 기사에 대한 코멘트를 발행하는 비-전문적 유저(소위, 웹 유저)를 포함할 수 있다.

[0044] 발명의 변형으로, 멀티미디어 콘텐츠가, 예를 들어, 적응된 콘텐츠 구축자에 의해 자동으로 생성될 수 있다는 점에 또한 유의해야 한다.

[0045] 각각의 크리에이터는 예를 들어, 크리에이터의 프로파일로부터 획득되거나 또는 이들이 생성한 콘텐츠에 관련된 특징들의 세트에 의해 기술될 수 있다. 예시적인 예로서, 이러한 특징은 생성된 콘텐츠의 수, 스타일리스트 언어, 이들의 콘텐츠에 언급된 명명된 엔티티의 중앙 수, 이들이 코멘트한 영화의 수 등에 대응할 수 있다.

[0046] 도 1은 예를 들어, 영화에 관련된 웹 커뮤니티에 속하는 주어진 콘텐츠 크리에이터(204)에 의해 생성된 고려된 멀티미디어 콘텐츠(예컨대 웹 콘텐츠)를 인리칭(enriching)하기 위한 방법 M의 주요 단계를 예시한다. 복수의 크리에이터(202, 204, 206, 208, 212, 214, 222, 224, 226) - 상기 웹 커뮤니티를 형성하는 - 가 도 2에 표현된다. 그것은 예를 들어, 전문 비평가 및 비전문가 멤버를 포함할 수 있다.

[0047] 도 1 및 2에 예시된 바와 같이, 방법 M은 크리에이터들의 몇몇 클러스터들을 설정하기 위해 복수의 크리에이터들의 클러스터링 단계(100)를 포함한다. 예를 들어, 도 2에 나타난 바와 같이, 크리에이터들(202, 204, 206, 208, 212, 214, 222, 224, 226)은 이들의 특징 세트에 따라 상이한 클러스터(200, 210, 220)에서 클러스터링될 수 있다. 각각의 클러스터는 계산된 유사성의 기준에 따라 특징들의 세트가 가까운 값을 획득하는 크리에이터들을 포함한다.



[0048] 클러스터링 단계(100)는 상기 방법 M의 후속하는 단계 동안 이루어질 작업량을 감소시키기 위한(특히, 방법 M을 구현하는 디바이스에서의 속도 및 메모리 소모의 견지에서) 사전-필터링 단계로서 작용한다.

[0049] 예를 들어, 클러스터링 단계(100)는 주어진 클러스터의 고려된 멀티미디어 콘텐츠가 인리칭될 때마다, 또는 복수의 크리에이터의 생성시에만, 또는 크리에이터가 상기 복수의 크리에이터에 추가될 때마다, 또는 때때로(예를 들어, 주기적 기반으로) 또는 크리에이터의 수의 비의 증가에 따라 수행될 수 있다.

[0050] 바람직한 실시예의 예시적인 그러나 비제한적인 예에 따르면, 크리에이터(202, 204, 206, 208, 212, 214, 222, 224, 226)는 K-평균 알고리즘을 사용하여 클러스터링된다. k-평균 클러스터링은 n개 관측의 세트( $x_1, x_2, \dots, x_n$ )를 구획하는 것을 목적으로 하고, 여기서 각각의 관측은 클러스터-내 제곱의 합산(WCSS)을 최소화하기 위해 k개 클러스터( $k \leq n$ )  $S = \{S_1, S_2, \dots, S_k\}$  내로의 d-차원 실수 벡터이다:

$$\arg \min_S \sum_{i=1}^k \sum_{x_j \in S_i} \|x_j - \mu_i\|^2$$

[0051]

[0052] 여기서,  $\mu_i$ 는  $S_i$ 에서의 포인트들의 평균이다.

[0053] 각각의 크리에이터는 관측  $x_i$ 에 의해 표현되고, (특징 세트에 대응하는) 자신의 특징 값들의 벡터에 따라 클러스터링된다. 주어진 크리에이터의 세트에 대한 클러스터의 수는 달라질 수 있다. 특히, 알고리즘(예컨대, Hartigan 알고리즘 또는 Krzanowski & Lai 알고리즘)이 k의 최적 값을 정의하기 위해 사용될 수 있다.

[0054] 이들 크리에이터 각각에 대해, 크리에이터를 기술하는 다수의 특징  $\{f_1, f_2, \dots, f_n\}$ 이 특징의 벡터를 형성하기 위해 정의된다.

[0055] 도 2에 도시된 바와 같이, k-평균 알고리즘 - 크리에이터(202, 204, 206, 208, 212, 214, 222, 224, 226)의 특징 벡터에서 사용된 - 은 크리에이터(202, 204, 206, 208, 212, 214, 222, 224, 226)의 몇몇 클러스터(200, 210 및 220)로의 구획을 초래한다.

[0056] 명백하게, 클러스터링 단계(100)는 예를 들어, 계층적 클러스터링 또는 스펙트럼 클러스터링과 같은 다른 클러스터링 기법들을 구현할 수 있다.

[0057] 클러스터링 단계(100)는 추가 데이터를 가지고 인리칭하기 위해 고려된 멀티미디어 콘텐츠의 주어진 크리에이터(204)에 대해 가장 관심 있는 크리에이터들의 사전-필터링으로서 작용한다. 그것은 이후, 상기 주어진 멀티미디어 콘텐츠의 크리에이터(204)와의 유사성을 가지는 관련 크리에이터(202, 206, 208)를 식별하는 것을 초래한다.

[0058] 또한, 클러스터(200)에 속하는 고려된 멀티미디어 콘텐츠의 주어진 크리에이터(204)를 고려함으로써, 방법 M은 다음 단계를 더 포함한다:

[0059] - 클러스터(200)의 각각의 관련 크리에이터(202, 206, 208)에 관련성의 스코어(232, 236, 238)를 할당하는 단계(단계 110)(일 변경으로, 클러스터의 관련 크리에이터들의 일부만이 관련성 스코어를 할당받았을 수 있다). 특히, 각각의 관련성 스코어는, 주어진 크리에이터(204)와 평가된 관련 크리에이터(202, 206, 208) 사이의 유사성의 기준에 따라, 주어진 크리에이터(204)와 연관되어 평가된다. 예로서, 유사성의 기준은 거리 기준일 수 있다. 거리는 클러스터(200)의 고려된 관련 크리에이터(202, 206, 208)의 특징 세트(세트의 적어도 하나 또는 몇몇 특징)를 주어진 크리에이터(204)의 특징 세트와 비교함으로써 추정될 수 있다. 실제로, 크리에이터들은 이들의 연관된 평가된 특징들의 근접도에 따라 클러스터 내에서 함께 그룹화된다. 특징들의 정확한 동일한 세트를 특징으로 하는 모든 크리에이터, 이들 값들의 세트에 인가된 임의의 종류의 거리 측정, 예를 들어, 유클리드 거리, 맨하탄 거리, 2개의 특징 값 사이의 최소 또는 최대 차이 등이 사용될 수 있다:

[0060] 유클리드 거리:  $dist_{euclidian} = \sqrt{\sum_{i < n} (x_i - x'_i)^2}$

[0061] 맨하탄 거리:  $dist_{manhattan} = \sum_{i < n} |x_i - x'_i|$

[0062] 최대 차이:  $dist_{max} = \arg \max_{i < n} (|x_i - x'_i|)$



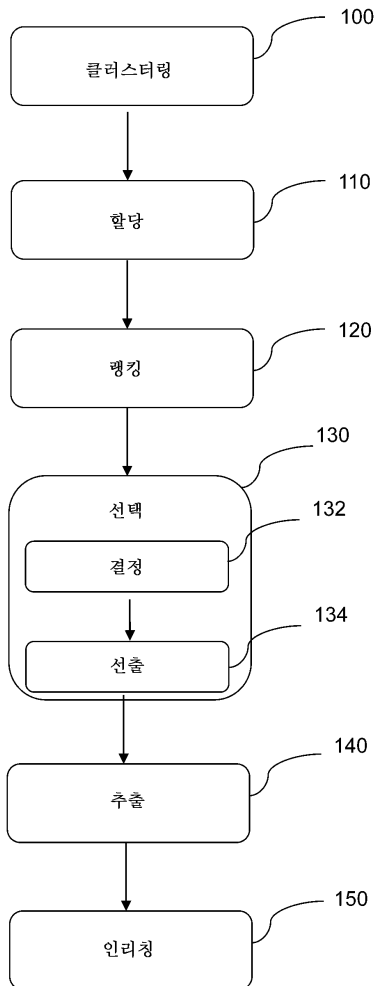
- [0063] 여기서,  $x_i$ (각각  $x'_i$ )는  $n$ 개 특징들 사이의 크리에이터  $x$ (각각 크리에이터  $x'$ )에 대한 수치 특징 값이다.
- [0064] - 그들의 관련성의 스코어(232, 236, 238)에 따라, 클러스터(204)의 관련 크리에이터(202, 206, 208)를 랭킹시키는 단계(단계 120);
- [0065] - 클러스터(200)의 랭킹된 관련 크리에이터(202, 206, 208)에 의해 전달된 일부 멀티미디어 콘텐츠 중에서, 주어진 크리에이터(204)의 고려된 멀티미디어 콘텐츠에 관련된 하나 이상의 추가적인 멀티미디어 콘텐츠를 선택하는 단계(단계 130). 추가적인 멀티미디어 콘텐츠가 전용 웹사이트로부터 이전에 크롤링(crawl)되거나 검색될 수 있으며, 주어진 멀티미디어 콘텐츠와 추가적인 멀티미디어 사이의 연관이 바람직하게는 크롤링 시간에 이미 공지되어 있다는 점에 유의해야 한다;
- [0066] - 선택된 추가적인 멀티미디어 콘텐츠(들)로부터 엘리먼트를 추출하는 단계(단계 140). 특히, 이러한 추출 단계는 자연어(natural language) 프로세싱 및/또는 텍스트-마이닝 기법에 기초할 수 있다. 이러한 추출 단계(140)는 (예를 들어, 하기에 기술되어 있는 바와 같이, 선출 서브-단계(132) 동안) 선택 단계(130)와 결합되고 병합될 수 있다;
- [0067] - 단계(130)에서 선택된 추가적인 콘텐츠에 기초하여 고려된 멀티미디어 콘텐츠를 인리칭하는 단계(단계 150). 특히, 상기 고려된 멀티미디어 콘텐츠는 메타데이터(예컨대, (추출된 엘리먼트로서) 추가 멀티미디어 콘텐츠의 일부분, 이 추가 콘텐츠에 대한 링크 등)의 포함에 의해 그리고/또는 고려된 멀티미디어 콘텐츠에 연관된 콘텐츠 디스크립터 내의 메타데이터, 예를 들어, 선택된 추가 멀티미디어 콘텐츠(들)로부터 추출된 텍스트 데이터의 추가에 의해 인리칭될 수 있다.
- [0068] 특히, 선택 단계(130)는 다음 서브-단계들을 포함한다:
- [0069] - 클러스터(200)의 랭킹된 크리에이터들 중, 가장 많이 관련된 크리에이터(들)를 결정하는 단계(단계 132). 결정 단계(132)는:
- [0070] ● 오직 하나의 랭킹된 크리에이터, 예를 들어, 클러스터(200)에 할당된 모든 관련성 스코어 중 가장 높은 관련성 스코어를 가지는 크리에이터; 또는
- [0071] ● "p"의 가장 높은 관련성 스코어를 가지는 몇몇 랭킹된 크리에이터("p"는 미리 결정된 수임); 또는
- [0072] ● 미리 정의된 관련성 임계치보다 더 큰 관련성 스코어를 가지는 몇몇 랭킹된 크리에이터
- [0073] 의 선택을 초래할 수 있다.
- [0074] - 결정된 관련 크리에이터에 의해 전달된 콘텐츠 중, 적어도 하나의 추가적인 멀티미디어 콘텐츠를 선출하는 단계(단계 134). 선출 서브-단계(134)는 예를 들어, 추가적인 멀티미디어 콘텐츠에 연관된 메타데이터에, 또는 상기 추가적인 콘텐츠 또는 (예를 들어, 추출 단계(140) 동안) 상기 추가적인 콘텐츠로부터 추출된 엘리먼트에 연관된 코멘트에 기초할 수 있다. 다른 변형에서, 선출 단계는 더 최근의 콘텐츠 또는 더 많이 언급된 콘텐츠 등의 선택을 초래할 수 있다.
- [0075] 다시 말해, 선출 서브-단계(134)는:
- [0076] - 가장 높은 관련성 스코어를 가진 랭킹된 크리에이터에 의해 전달된 하나 이상의 추가적인 멀티미디어 콘텐츠;
- [0077] - "p"의 가장 높은 관련성 스코어를 가진 랭킹된 크리에이터에 의해 전달된 콘텐츠 중에서, 또는 특정한 미리 결정된 임계치보다 더 큰 관련성 스코어를 가지는 랭킹된 크리에이터에 의해 전달된 콘텐츠 중에서, "m"개의 가장 유사한 콘텐츠와 같은 몇몇 추가적인 멀티미디어 콘텐츠("m"도 또한 미리 결정된 수임)
- [0078] 의 선택을 초래할 수 있다.
- [0079] 예로서, 그러한 선택된 추가적인 멀티미디어 콘텐츠는 이후 - 인리칭 단계(150) 동안 - 고려된 멀티미디어 콘텐츠 그 자체에, 또는 상기 고려된 콘텐츠와 연관된 디스크립터에 추가될 수 있다.
- [0080] 변형으로, 방법 M의 단계가 명백하게는 이전에 기술된 것과는 상이한 순서로 수행될 수 있다는 점이 이해되어야 한다.
- [0081] 발명에 따르면, 방법 M은 바람직하게는, 고려된 멀티미디어 콘텐츠의 크리에이터로부터 또는 콘텐츠 뷰어로부터의 소수의 동작을 가지고(또는 심지어 어떤 동작도 없이) 다소 자동으로 수행될 수 있다.

- [0082] 추가로, 발명으로 인해, 크리에이터의 연관성은 관련 크리에이터 중 적어도 하나에 의해 발행된 콘텐츠의 관련성(단계 110, 120, 130)에 추가하여 고려될 수 있다(단계 100). 이는 가장 많이 관련된 크리에이터의 가장 정확한 콘텐츠의 선택을 초래할 수 있다. 주어진 멀티미디어 콘텐츠와 연관된 메타데이터의 품질은, 소수의 유저 상호작용을 가지고(또는 심지어 유저 상호작용 없이) 개선될 수 있다.
- [0083] 또한, 이전에 기술된 멀티미디어 콘텐츠를 인리칭하기 위한 방법 M의 단계는, 예를 들어, 인터넷 타입 네트워크와 같은 통신 네트워크에 접속될 수 있는, 도 3에 도시된 바와 같은 디바이스(400)에 의해 구현될 수 있다.
- [0084] 특히, 고려된 멀티미디어 콘텐츠의 주어진 크리에이터(204)를 고려함으로써, 디바이스(400)는:
- [0085] - 둘 이상의 관련 크리에이터를 포함하는 클러스터를 전달하기 위해 복수의 크리에이터(202, 204, 206, 208, 212, 214, 222, 224, 226)를 클러스터링하도록 구성되는 클러스터링 모듈(112);
- [0086] - 상기 관련 크리에이터와 주어진 크리에이터(204) 사이의 유사성의 기준으로부터 결정된 관련성 스코어를, 주어진 크리에이터(204)를 포함하는 클러스터(200)의 각각의 관련 크리에이터(202, 206, 208)에 할당하도록 구성되는 할당 모듈(106);
- [0087] - 그들의 관련성 스코어에 따라 클러스터(200)의 관련 크리에이터(202, 206, 208)를 랭킹시키도록 구성되는 랭킹 모듈(104);
- [0088] - 복수의 멀티미디어 콘텐츠 중에서, 상기 클러스터(200)의 랭킹된 관련 크리에이터(202, 206, 208)에 의해 전달된, 고려된 멀티미디어 콘텐츠에 관련된 하나 이상의 추가 멀티미디어 콘텐츠를 선택하도록 구성되는 선택 모듈(131);
- [0089] - 예를 들어, 자연어 프로세싱으로 인해, 고려된 멀티미디어 콘텐츠에 관련된 적어도 하나의 엘리먼트를 상기 추가 멀티미디어 콘텐츠(들)로부터 추출하도록 적용되는 추출 모듈(133);
- [0090] - 예를 들어, 추출된 엘리먼트(들)로 인해, 상기 추가 멀티미디어 콘텐츠(들)에 기초하여 고려된 멀티미디어 콘텐츠를 인리칭하도록 구성되는 인리칭 모듈(108)
- [0091] 을 포함할 수 있다.
- [0092] 이들 모듈은 (도 3에 도시되지 않은) 프로세싱 유닛의 마이크로프로세서에 의해 구동될 수 있다.
- [0093] 당업자에 의해 이해될 바와 같이, 본 원리들의 양상은 방법, 디바이스, 컴퓨터 판독가능한 매체 또는 컴퓨터 프로그램 제품으로서 구현될 수 있다.
- [0094] 따라서, 본 원리들의 양상은 전적으로 하드웨어 실시예, 전적으로 소프트웨어 실시예(펌웨어, 상주 소프트웨어, 마이크로-코드 등을 포함함), 또는 모두가 일반적으로 본원에서 "회로", "모듈", 또는 "시스템"으로서 지칭될 수 있는 소프트웨어 양상과 하드웨어 양상을 결합한 실시예의 형태를 취할 수 있고, 그 전체는 단일 디바이스에서 또는 임의의 종류의 수단에 의해 함께 접속된 많은 디바이스들에서 실시된다. 또한, 본 원리들의 양상은 컴퓨터 판독가능한 저장 매체의 형태를 취할 수 있다. 하나 이상의 컴퓨터 판독가능한 저장 매체(들)의 임의의 결합이 활용될 수 있다.
- [0095] 기재, 청구항 및 도면에 개시된 인용물들은 독립적으로 또는 임의의 적절한 조합으로 제공될 수 있다. 특징은, 적절한 경우, 하드웨어, 소프트웨어 또는 그 둘의 조합에서 구현될 수 있다.
- [0096] 청구항에 나타나는 참조 부호는 오직 예시에 의한 것이며, 청구항의 범위에 대한 어떠한 제한적인 영향도 가지지 않을 것이다.
- [0097] 이 발명은 그 바람직한 실시예에서 기술되었으며, 당업자의 능력 내에서, 그리고 발명 능력의 실시 없이도 다수의 수정들 및 실시예들이 용이하다는 점이 명백하다. 따라서, 발명의 범위는 후속하는 청구항의 범위에 의해 정의된다.
- [0098] 발명의 청구항에서, 특정된 기능을 수행하기 위한 수단으로서 표현된 임의의 엘리먼트(예를 들어, 클러스터링 모듈(112), 할당 모듈(106), 랭킹 모듈(104), 선택 모듈(131), 추출 모듈(133), 인리칭 모듈(108) 등)는 예를 들어, a) 그 기능을 수행하는 회로 엘리먼트(예를 들어, 하나 이상의 프로세서)의 결합, 또는 b) 따라서, 기능을 수행하기 위해 그 소프트웨어를 실행하기 위한 적절한 회로와 결합된, 펌웨어, 마이크로코드 등을 포함하는 임의의 형태의 소프트웨어를 포함하는 그 기능을 수행하는 임의의 방식을 포함한다. 이러한 청구항에 의해 정의된 본 원리는 다양한 인용된 수단에 의해 제공되는 기능성이 청구항이 요구하는 방식으로 함께 결합되어 온다

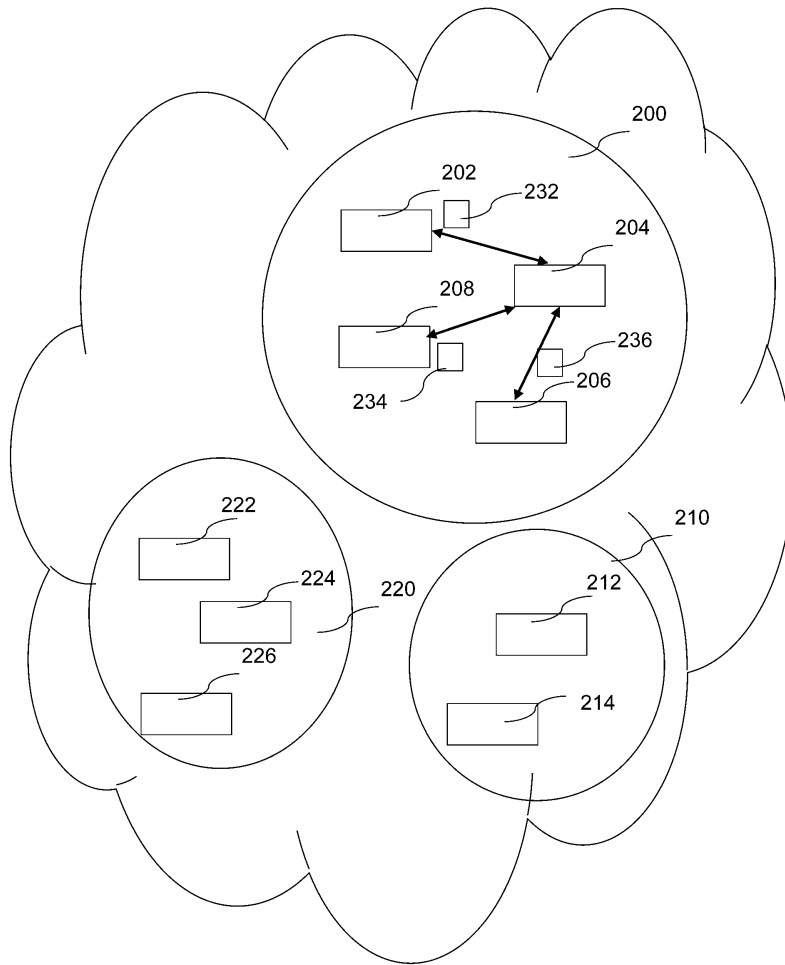
는 사실에 있다. 따라서, 그 기능성을 제공할 수 있는 임의의 수단은 본원에 도시된 것과 등가인 것으로 간주된다.

## 도면

### 도면1



도면2



도면3

